M3 · 第 8 讲

搭第一个智能体

以会议纪要为例，从零到可用

阅读约 35 分钟

练习约 30 分钟

产出：一个可运行的窄场景智能体

本节目录

开场：一个能跑通的智能体，比十个 demo 有用

本课目标

一、第一步：定输入输出

二、第二步：写系统提示

三、第三步：挂知识库（可选）

四、第四步：加工具

五、第五步：测试

六、发布：先灰度

七、30 分钟练习

八、五个常见坑

九、检查清单

十、延伸阅读

开场：一个能跑通的智能体，比十个 demo 有用

看别人演示智能体，感觉什么都能做；自己动手，第一版往往连输入格式都不对。原因很简单：演示挑的是最顺的例子，真实任务充满例外。

这一讲用一个窄场景——**会议纪要整理**——把搭建流程走完整。窄，才能跑通；跑通，才有第二版。

本课目标

- 用五步把一个场景做成可运行的智能体；
- 写一份能约束行为的系统提示；
- 用 5 条测试用例验收，而不是凭感觉；
- 知道第一版该砍掉什么。

一、第一步：定输入输出

先写死格式，不要边做边改。

项	定义
输入	一段会议录音转写文本（含说话人、时间戳）
输出	三段：结论、待办（谁/做什么/何时）、待确认问题
长度	结论不超过 5 条，待办不超过 8 条
失败	转写缺失或无人名时，输出“信息不足”并列出缺什么

输入输出不清，提示词写得再好也不稳定。

二、第二步：写系统提示

一个够用的模板：

> 你是会议纪要助手。你的任务是把转写文本整理成三段式纪要。 规则： 1. 只使用原文信息，不补充常识； 2. 每条待办必须包含负责人与时间，缺失就标“待确认”； 3. 结论必须是已达成的决定，不是讨论过程； 4. 输出固定三段，不加开场白； 5. 原文听不清或矛盾，列进“待确认问题”。

系统提示只写“规则”，不要把例子塞满。例子放少样本区，规则放这里。

三、第三步：挂知识库（可选）

如果你们有固定的项目名、人员表、术语表，挂进去能显著减少错名。

注意两件事：

- 知识库只放**稳定事实**，不要放会变的排期；
- 每份文档加标题与更新时间，方便检索命中。

四、第四步：加工具

第一版只加一个工具就够：**写入文档 / 发送到群**。不要一上来接五个系统。

加工具前问三句：

1. 失败了会不会造成外部影响?
2. 需不需要人工确认?
3. 能不能重试且不重复?

有外部影响的动作，一律加确认节点。

五、第五步：测试

准备 5 条用例，覆盖正常与异常：

编号	输入特征	期望
T1	标准 30 分钟会议	三段完整，待办有负责人
T2	有人没来，任务未定人	负责人标“待确认”

编号	输入特征	期望
T3	转写有大量错字	不臆测，列入待确认
T4	只有 3 分钟闲聊	输出“无明确结论”
T5	两个决定互相矛盾	两条都列出并标冲突

全部通过再发布。任何一条失败，先改规则，不要靠“再跑一次”。

六、发布：先灰度

- 先给自己用一周；
- 再给 2–3 个同事用，收集失败样本；
- 每周看一次日志：失败率、人工修改率、平均耗时；
- 稳定两周后再扩大范围。

七、30 分钟练习

1. 选一个窄场景（会议纪要/周报汇总/客服分类）（5 分钟）；
2. 写输入输出表（5 分钟）；
3. 写 5 条系统规则（10 分钟）；
4. 准备 3 条测试用例并跑一遍（10 分钟）。

八、五个常见坑

1. **场景太宽**：“帮我处理所有文档”无法验收；
2. **规则写成愿望**：“尽量准确”不可执行；
3. **没有失败输出**：模型会硬编一个答案；
4. **工具权限过大**：第一版就能删数据；
5. **只看成功案例**：失败样本才是迭代依据。

九、检查清单

- [ ] 输入输出格式写死；
- [ ] 系统提示只有规则，没有废话；
- [ ] 至少 3 条异常用例通过；
- [ ] 外部动作有人工确认；
- [ ] 有日志可回看。

十、延伸阅读

- 知识库 1.3 AI Agents（智能体）
- 知识库 3.2 工具：辅助学习工具和资源
- 知识库 4.3 团队搭建：AI 招聘/组队

> 下一讲：知识库问答与工作流。你会把十份文档变成一个能答对问题的助手，并知道答错时该查哪里。